University of Zurich[UZH]

Zurich Open Repository and Archive
University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2019

Hearing the silence: finding the middle ground in the spatial humanities? Extracting and comparing perceived silence and tranquillity in the English Lake District

Chesnokova, Olga ; Taylor, Joanna E ; Gregory, Ian N ; Purves, Ross S

Abstract: We analyse silence and tranquillity in historical and contemporary corpora to understand ways landscapes were—and are—perceived in the Lake District National Park in England. Through macro and microreading we develop a taxonomy of aural experiences, and explore how changes to categories of silence from our taxonomy—for instance, the overall decline in mentions of absolute silence—provide clues to changes in the landscape and soundscape of the Lake District. Modern authors often contrast silence with anthropogenic sounds, while historical authors adhere to a cultural construction where the Lake District is presented as a tranquil area by ignoring industrial sounds. Using sentiment analysis we show that silence and tranquil sounds in our corpora are, as a whole, more positively associated than random text from the corpora, with this difference being especially marked in contemporary descriptions. Focusing closely on individual texts allows us to illustrate how this increased positivity can be related to the emergence of silence and tranquillity as valuable components of landscape. Mapping our corpora confirmed the influence of Wordsworth's writing on descriptions of silence; and revealed the co-location of pockets of tranquillity near to transport arteries in contemporary descriptions.

DOI: https://doi.org/10.1080/13658816.2018.1552789

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: https://doi.org/10.5167/uzh-159816
Journal Article
Published Version

The following work is licensed under a Creative Commons: Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.

Originally published at:
Chesnokova, Olga; Taylor, Joanna E; Gregory, Ian N; Purves, Ross S (2019). Hearing the silence: finding the middle ground in the spatial humanities? Extracting and comparing perceived silence and tranquillity in the English Lake District. International Journal of Geographical Information Science, 33(12):2430-2454.
DOI: https://doi.org/10.1080/13658816.2018.1552789

INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE
https://doi.org/10.1080/13658816.2018.1552789

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Hearing the silence: finding the middle ground in the spatial humanities? Extracting and comparing perceived silence and tranquillity in the English Lake District

Olga Chesnokova [a], Joanna E. Taylor [b], Ian N. Gregory[c] and Ross S. Purves[a]

[a]Department of Geography, University of Zurich, Zurich, Switzerland; [b]Presidential Academic Fellow in Digital Humanities, University of Manchester, Manchester, UK; [c]Department of History, Lancaster University, Lancaster, UK

ABSTRACT

We analyse silence and tranquillity in historical and contemporary corpora to understand ways landscapes were—and are—perceived in the Lake District National Park in England. Through macro and microreading we develop a taxonomy of aural experiences, and explore how changes to categories of silence from our taxonomy—for instance, the overall decline in mentions of absolute silence—provide clues to changes in the landscape and soundscape of the Lake District. Modern authors often contrast silence with anthropogenic sounds, while historical authors adhere to a cultural construction where the Lake District is presented as a tranquil area by ignoring industrial sounds. Using sentiment analysis we show that silence and tranquil sounds in our corpora are, as a whole, more positively associated than random text from the corpora, with this difference being especially marked in contemporary descriptions. Focusing closely on individual texts allows us to illustrate how this increased positivity can be related to the emergence of silence and tranquillity as valuable components of landscape. Mapping our corpora confirmed the influence of Wordsworth's writing on descriptions of silence; and revealed the co-location of pockets of tranquillity near to transport arteries in contemporary descriptions.

ARTICLE HISTORY
Received 4 May 2018
Accepted 22 November 2018

KEYWORDS
Landscape; tranquillity; temporal changes; text analysis

1. Introduction

Human perception is shaped by our senses: the ways in which we experience the world are driven by what we see, hear, smell, touch and taste. The importance of this multi-sensory perception of landscape is emphasised in policy documents; the European Landscape Convention, for instance, defines landscapes as 'an area, as perceived by people, whose character is the result of the action and interaction of natural and/or human factors' (Council of Europe 2000, p. 2). A diverse range of academic fields—including human geography, landscape ecology, history and computer science—have increasingly recognised the need to move away from considering only what we see, to other ways of perceiving landscapes (e.g. Smith 1994, Pijanowski *et al*. 2011, Quercia and Schifanella 2015). This paper joins this

CONTACT Olga Chesnokova olga.chesnokova@geo.uzh.ch

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

new tradition of scholarship. From the perspective of spatial computing for the digital humanities, we will demonstrate how blending research methods from literary studies, corpus linguistics and geographical information science—with inputs from disciplines including geography and history—can offer fresh perspectives on landscape change with consequences for the ways we understand a location's development and its social, cultural and ecological status. As we will show, studying landscape(s) is an inherently interdisciplinary endeavour.

In purely spatial terms, landscapes matter because they are heterogeneous: different characteristics make different landscapes special, and drive the desire to protect unique or exemplary landscapes (Tudor 2014). Practically speaking, landscape characterisation based on perception typically combines existing spatial data and implicitly links this to—usually visual—perception. The senses which are thus modelled are always, by definition, incomplete. First, they reflect a particular set of cultures and practices where specific elements are considered more important in, for instance, planning decisions. Second, our ability to conceptualise, abstract and model spatially variable processes in technologies like GIS is limited. Some perceived elements of landscapes are, at least over short intervals of time, easier to conceptualise and model by knowing something about the constituent physical make-up of a scene: at their simplest, perceivable boundaries in a landscape indicate land use or land cover (Turner 2006); viewsheds (e.g. Lake *et al*. 1998, Fisher *et al*. 2004) provide a route to landscape vistas; and coherence can be modelled through shape and distribution of co-occurring parcels (e.g. McGarigal and Marks 1994). Such approaches provide a starting point, at least, for generating plausible spatial models that reflect what might be sensed at some locations. However, landscapes are not static; they are subject to both natural and anthropogenic processes which in turn lead to change.

Understanding and documenting landscape change is an important task in landscape planning, since it provides a baseline with respect to both objective and subjective notions of landscape over time. Despite a recognition that GIS can capture change—either through snapshots or process-based models (Grenon and Smith 2004)—in reality models of landscape change used for practical purposes focus almost exclusively on snapshots at particular moments in time. Even where we take an apparently simple approach to measuring change in the recent past, multiple challenges arise: changes in sensors and their capabilities (Sexton *et al*. 2016); changes in ontologies (Comber *et al*. 2005); cultural and linguistic differences (Burenhult and Levinson 2008); and changes in computational methods and representations (Vasconcelos *et al*. 2002) all affect landscapes over time. As we go further back in time to map historical landscapes, these challenges become more pronounced; primary data describing physical landscapes become increasingly scarce and the complexity of relating historical representations to current data models increases.

Nonetheless, such approaches typically assume that some basic spatial data exist, such as in the form of cartographic products (Leyk *et al*. 2006, Fuchs *et al*. 2015), aerial photographs or satellite imagery (Van Den Berghe *et al*. 2018). Historical Landscape Characterisation—a technique pioneered in England that seeks to model and interpret landscapes with respect to their historical development (Turner 2006, Fairclough and Herring 2016)—is an example of such an approach. Here, the importance of perception in understanding and representing past landscapes in mappable forms is clear. Yet, this process of interpretation is carried out by experts, and has thus been criticised in the context of Landscape Character Assessment as being dominated by values attributed

through 'objective' outsiders (Butler 2016). This criticism, and the relating tension between what are viewed as positivist or oversimplified mappings of landscapes on the one hand, and a lack of pragmatism and contribution to real societal needs to describe and monitor change on the other, has multiple parallels with well-known debates in Geographic Information Science (Pickles 1995, Rundstrom 1995).

We do not propose to rehash these debates here, but rather to try to find some middle ground. As a starting point, we understand written accounts as a window to perception. We use a blend of interdisciplinary methods to analyse text and thus show how landscape is—and was—perceived. We focus on one particular landscape: the English Lake District, a UNESCO World Heritage Site in the North West of England, and one of the most comprehensively recorded landscapes in the world (Nomination 2015). Indeed, the UNESCO designation recognises the region as a 'cultural landscape', so significant has been the effect of art and literature on its historical and contemporary character.

More specifically, what we are interested in here is the perception of a particular rural soundscape. Written sources are particularly important for providing a route to understanding perceptions of experiences which are, otherwise, ephemeral. We explore what written accounts of personal experiences might reveal about the role of sound and, in particular, silence in response to (perceived) natural landscapes. We will see that texts preserve, however imperfectly, sounds which risk being forgotten as social, cultural and technological contexts change (Lowenthal 1976, Smith 1994, p. 233). Although written descriptions do not necessarily reflect the objective soundscape, they do offer insights into personal experiences of place. More than that, they indicate something of the changing social and cultural status of sounds. As we demonstrate, written accounts can provide us with one—admittedly incomplete, yet nevertheless significant—way of understanding what people describe when communicating about both contemporary and historical landscapes, and why this matters for the development of these locations. If these descriptions are about the same places, at different times, then we can pose an important question: can we use written accounts to characterise changes in perceived sounds and silences in landscapes across both space and time?

To answer this question, we first explore how sound and soundscapes have emerged as important components of landscape studies. We then discuss the emergence of notions of tranquillity and quiet at the turn of the eighteenth century, and show how these ideas continue to influence modern-day values associated with rural peace, silence and tranquillity. We introduce our study area, the Lake District, and the two corpora on which our study is based. Finally, we explain the interdisciplinary approach to text analysis which has allowed us to offer conclusions about the nature of tranquillity, and to understand changes to the ways in which the Lake District landscape is valued.

2. Background

2.1. *Sound and silence*

Sound affects us more consistently, perhaps, than any other sense; as Bruce R. Smith observes, we are 'surrounded – and filled – by a continuous field of sound' (1999, p. 9). Yet, it is only comparatively recently that scholarship has begun to reflect on the importance of multisensory perception to human understandings of place and space.

The new awareness of sound represented by the work of scholars like Susan J. Smith, Mark Smith and Alain Corbin (Corbin 1986, Smith 1994, 2004) is indebted to R. Murray Schafer's work on the World Soundscape Project in the 1970s and 1980s. Schafer was predominantly interested in the components which made up an area's soundscape, but more recent work has moved on from the 'acoustic ecology' promoted by the World Soundscape Project (Truax 1978), preferring instead to highlight a 'soundscape ecology' that focuses on interactions between different acoustic elements and their environment (Pijanowski *et al*. 2011). This work delivers new assessments of the ways that humans interact with—or are affected by—the soundscapes they encounter. Specifically, as Pijanowski and his co-authors suggest, more research is needed into 'how natural sounds influence the development of individuals' sense of place, place attachment, or connection to nature', as well as the factors which 'affect human (in)tolerance of soundscape changes, especially where those changes increase noise' (2011, p. 209).

As these scholars imply, acoustic experiences are as subject to contemporary fashion as visual ones. Peter Coates puts it neatly when he writes that '[j]ust as beauty is in the eye of the beholder, noise frequently resides in the ear of the listener' (Coates 2005, p. 641). In short, what is interpreted as a sound—rather than as noise—changes over time. As Isabelle Bour suggests, such transitions can occur slowly over centuries, or as quickly as the course of a day: 'a buskers trumpet,' she concludes, 'is likely to be perceived as sound at midday but as noise at midnight' (Bour 2016). The way a sound is interpreted depends on the characteristics of the sound, including its volume, as well as the affective response it initiates in the listener (MacFarlane *et al*. 2004, p. 134).

The late eighteenth century catalysed this emergence of listening as what Sophia Rosenfeld calls 'a cultural effect as much as a physiological one' (Rosenfeld 2011, p. 318). Debates about what constituted sound, and what noise, emerged at this moment—from when the earliest texts on which we focus here date—when philosophers began developing an acoustic aesthetics that paralleled the development of the picturesque for visual phenomena (Agnew 2012, Joy 2014, Donaldson *et al*. 2017). Thinkers such as William Duff and James Beattie agreed that the combination of certain sounds formed inherent harmonies, and so might produce pleasurable emotional responses in the listener (Dubois 2016). Others, though, grated on the listener; to describe an acoustic experience as 'noisy' has always been pejorative. Paul Hegarty explains that noise is negative because it 'can never be positively, definitely and timelessly located' because it is emblematic of elements that society wishes to resist (Hegarty 2007, p. ix).

This emergence of sound and noise as important topics for discussion is perhaps unsurprising in an age that witnessed profound changes to its soundscapes. Noise control was a prominent concern, in urban environments especially, throughout the eighteenth century. By the Victorian period, city soundscapes had intruded into the countryside: the railway screeched and rattled over the new lines that criss-crossed the country; cacophonies from emerging industrial centres echoed around the surrounding area; and, in towns and cities, new forms of making sound and noise contributed to a sonic shift with profound cultural consequences (Picker 2003, p. 5). Even before the motor car's arrival, the nuances of the soundscape—and, particularly, the delicacies of natural sounds—had been largely masked by human noises. The result was a shrinking of what Bruce Smith calls the 'acoustic horizon'; modern sounds, especially the low drone of the internal combustion engine, obscure other low frequency sounds and

dramatically reduce the distance at which all sounds can be heard by the human ear (1999, p. 51). Smith believes that quiet environments, away from the drone of traffic and hubbub of daily life, expand our acoustic horizons and enable deeper, more meaningful connections between the self and the world (Smith 1999, p. 74).

By the turn of the eighteenth century, urban residents in particular began to long for respite from the constant din of the modern world, and to seek out locations that still maintained—at least relatively—nuanced, natural soundscapes. Following the rapid expansion of industrialisation towards the end of the eighteenth and into the beginning of the nineteenth centuries, it did not take long for cultural evaluations of quietness, as well as noise and sound, to shift. Before then, quietness had seemed to indicate a lack of civilisation; it was a marker of what John Fisher calls 'untrammeled nature', and the impulse was to tame it (1999, p. 27). But as towns and cities grew noisier, those who had the financial means began to seek silence. By the mid-nineteenth century, 'quiet' became equated with 'peace' (Ammer 2013), and the remoteness offered by less accessible regions—such as the Lake District—began to be valued as much for the respite they could offer from the urban din as for their picturesque beauty.

As we will see, contemporary landscape preferences maintain this need for peace and quiet; tranquillity requires sounds and silences that generate states of repose. The search for peace and quiet away from urban centres is not merely a desire for a break from day-to-day pressures; it is also a necessary part of the human requirement for connection with the natural world, and for related feelings of calm. Indeed, one of the main reasons people today give for visiting rural landscapes is the search for peace and quiet; in a 2001 survey conducted by the UK's Department for Environment, Food and Rural Affairs, 'tranquillity' was cited by 58% of respondents as their main motivation for spending time in the countryside, ahead of scenery (46%), open space (40%), fresh air (40%), or plants and wildlife (36%) (MacFarlane *et al*. 2004, p. 7). National organisations charged with the care of these landscapes have increasingly recognised the importance of sound, as well as sight, on visitor experiences (Miller 2008): in the case of the American National Park Service (NPS), for instance, the preservation of soundscapes was enshrined in an act signed by President Woodrow Wilson in 1916. Today, the NPS runs a comprehensive programme to promote the conservation of soundscapes in their care, and to educate their visitors about the importance of natural sounds as 'part of a web of resources [that are] vital to park ecosystems'.[1]

In Britain, the protection of the Lake District's tranquillity—and the 'sense of space and freedom' it engenders—was one of the motivating factors behind the Lake District National Park's successful application for UNESCO World Heritage Site status in 2017 (Nomination 2015, 2.72). This designation, the historical association of the region with peace and quiet, and the volume of writing it has inspired make it an exemplary location to test the potential of our approach.

3. Writing Lakeland: the corpus of Lake District writing and Geograph

The boundaries of the Lake District National Park (Figure 1) were established in the mid twentieth century, but the region has been celebrated for its picturesque beauty, opportunities for outdoor activities (including mountain climbing and wild swimming) and comparative remoteness since the late eighteenth century. It is particularly

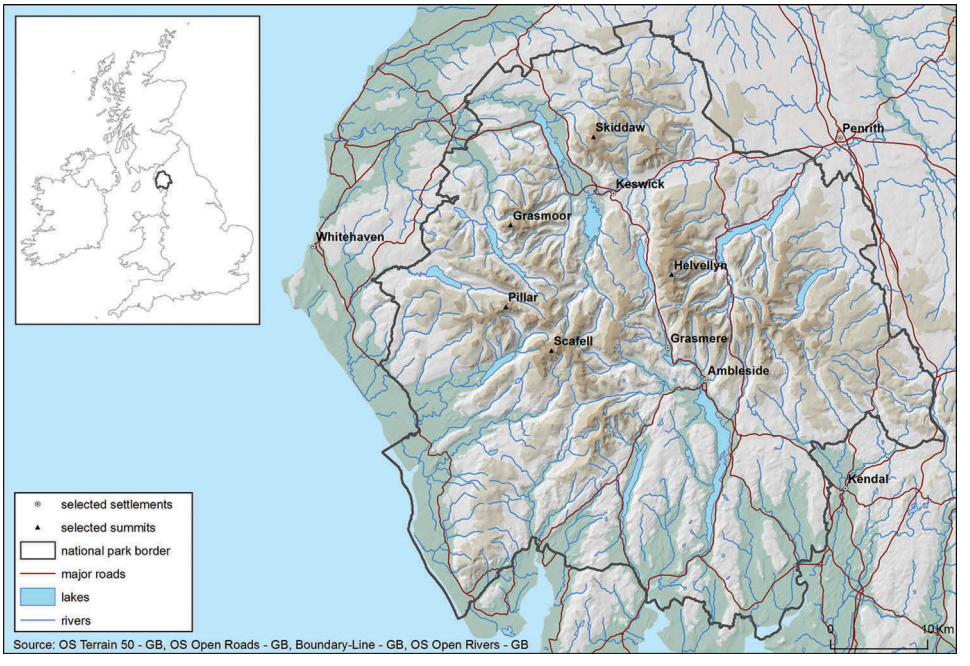

Figure 1. Our study area: The Lake District in North West England.

renowned for the authors and artists who have made their homes there, or found inspiration in its dramatic upland landscapes: William and Dorothy Wordsworth, John Ruskin, Beatrix Potter and Arthur Ransome all lived in the area, and their writing and wider agricultural, geological and artistic interests have—in different ways—shaped today's Lake District. Alongside these famous figures, a host of lesser-known writers and the increasingly multitudinous voices of visitors to the region have influenced the conservation practices that maintain this area as a particular kind of cultural landscape (Nomination 2015, Donaldson *et al*. 2017).

We are especially interested in written responses to the Lake District from the eighteenth century to the present day extracted from two sources: the historical Corpus of Lake District Writing (CLDW)[2] and the contemporary Geograph Project.[3] The Corpus of Lake District Writing is a georeferenced collection of writing about the Lake District and the surrounding area; broadly speaking, it is interested in the modern county of Cumbria, which was formed in the 1970s by merging the old counties of Cumberland and Westmorland with the Lancashire Hundred of Lonsdale North of the Sands. Currently, it comprises 1.5 million words, consisting of 80 texts that include novels, poetry, epistolary fiction, non-fictional essays, topographical accounts and travel writing about the region (Donaldson *et al*. 2015, Gregory and Donaldson 2016, Murrieta-Flores *et al*. 2017). To focus as closely as possible on lived experiences of the historical Lake District's soundscape, for this project we selected only the non-fictional prose accounts as the most similar to our second corpus. The version of the CLDW we discuss in this article, then, contains 61 texts written by 55 authors, and comprises 1.3 million words.

Our second corpus is constituted of text from the crowdsourced project Geograph British Isles. Geograph, launched in 2005, aims to document landscape at the scale of

1km grid squares. The site combines photographs with textual descriptions, written by the photographers themselves, and has an active community of more than 12,000 users. Like most User Generated Content platforms, many of Geograph's contributions tend to be made by a small number of users, most of whom are men over the age of 50.[4] For the Lake District and its immediate surroundings, there are 90,705 photographs, contributed by 1218 authors. Of these, 64,795 include descriptions of photographs, which total 1.4 million words written by 1076 authors. Geograph is well suited to our task since, firstly—and unlike other photo-sharing platforms like Flickr—images are associated with full text descriptions rather than tags which are better suited to semantic analysis. Secondly, geographic coverage is not strongly biased to urban areas because of the use of a points system that motivates documentation of new grid squares (Antoniou *et al*. 2010). Thirdly, Geograph explicitly focuses on collecting descriptions of experienced (and thus perceived) landscapes (Chesnokova and Purves 2018). Example extracts from both the CLDW and Geograph can be found in Figures 5 and 6.

Both corpora are composed by authors who, by the nature of the genres in which they write, are highly aware of the relationship between location and description. Comparing these sources therefore allowed us to assess a diachronic sweep of writing about the region, and to trace in it evolving understandings of the Lake District's status as an enclave for tranquillity in an increasingly industrialised, mechanised and digitised world. Nonetheless, there are differences between the two corpora. Firstly, and most importantly, the texts in each were written with very different aims and audiences in mind. Nevertheless, both corpora contain references to first-person perceptions of silence, and document something of the lived experience of listening to the landscape. Secondly, Geograph has significantly more individual authors, and thus has the potential for a wider range of perspectives (although, as we will see shortly, their approaches are remarkably homogeneous). Thirdly, the sentence length in the CLDW is—unsurprisingly, given the dates of composition of the texts—almost twice as long as in the modern corpus, and these are much more heavily punctuated. Nevertheless, with 43,269 sentences (for the CLDW) and 98,206 (for Geograph) sentences each, our corpora are of a comparable size.

4. Finding the middle ground

4.1. *Overview of the process*

To identify, compare and analyse text relating to sounds in these two corpora, we combined so-called macro and microanalysis of the texts (Jockers 2013). We iteratively applied a mixture of text analysis methods, guided and informed by the contemporary acoustic-cultural contexts in which the texts were written, to process the works quantitatively and to identify features for further exploration. By also qualitatively microanalysing the texts in our corpora, we were able to ask more specific questions about the relationship between personal perceptions of landscape and soundscape, and to develop a more detailed understanding of how the soundscape affects individuals' senses of peace and tranquillity in the Lake District National Park. Finally, we mapped a subset of texts from Geograph and the CLDW to explore spatial patterns in the extracted texts.

Our approach can be grouped into four main tasks, listed here in the order undertaken in our methodological pipeline (thus, e.g., annotation was carried out after an initial exploration of the corpora):

- Preprocessing: a set of basic methods to prepare a text snippet for analysis.
- Search and comparison: extraction of texts potentially describing sounds and comparison of their associated content within and between corpora.
- Annotation: classifications of the nature of silence and its emitters in text snippets.
- Enhancement: the calculation of additional properties related to text snippets, for example in our case the sentiment associated with a given snippet and the association of individual texts with locations.

The first of these steps, preprocessing, is more or less independent of the analysis which follows (though the choices made here may influence our results). Typical pipelines initially chunk corpora into documents, and then divide documents (in this case, individual accounts from the CLDW or photographs from Geograph) into sentences and tokens (tokens are the finest units of analysis with which we concern ourselves, and can include words, punctuation and other processed elements of text such as lemmas or stems (Manning and Schutze 1999)). In the case of Geograph, extracting sentences is a relatively simple task, since the texts largely consist of short captions for the associated image. The CLDW, on the other hand, presented more challenges, including idiosyncratic punctuation, case and hyphenation (Butler *et al*. 2017, Donaldson *et al*. 2017). Nonetheless, in both cases sentence extraction and tokenisation were carried out using the NLTK Python Library with no modifications.[5] Having tokenised the texts, we carried out part of speech tagging, removed stop words and normalised all tokens to lower case.

4.2. *Seeking the silence*

An initial reading of our two corpora showed that both contained a wealth of descriptions relating to peaceful sounds in the Lake District, and our first task was to build two sub-corpora containing only text snippets that related to such acoustic experiences. As we have seen, over the course of the period in which we are interested here, interpretations of quietness shifted; it transitioned from being understood as a symptom of a lack of civilisation into a desirable characteristic for the experience of tranquillity. In light of this change, we used the Historical Thesaurus of English to generate a set of seed terms that are related to quiet sounds in the 'Inaudibility', 'Faintness/weakness' and 'Quietness/Tranquillity' categories. We removed terms which were only in use before 1750, but retained those which were either current in the eighteenth or nineteenth centuries, or which came into use prior to that date and have continued to be used in the present day. An initial search demonstrated that a subset of the terms thus selected were highly ambiguous in both the CLDW and Geograph extracts (e.g. *rest, sleep, dead*) and we removed these from the list. To further minimise the effects of word sense ambiguity, we used part of speech tagging to filter more terms; for instance, we retained *still* when used as an adjective (sense: 'not moving or not making a sound'), but removed it where it was employed as an adverb (sense: 'even now'). It is important to note that

our aim in these searches was to maximise the recall of descriptions related to silence, while achieving reasonable precision.

The resulting sub-corpora consisted of sentences which were likely—though not guaranteed—to be related to quiet sounds or peaceful experiences. Table 1 shows some of these sub-corpora's basic properties in relation to the main corpora from which they were extracted. Interestingly, potential descriptions of silence are more or less ubiquitous in the CLDW (89% of authors referred to silence), but much rarer overall in Geograph (mentioned by only around 10% of authors). This difference might illustrate the dominance of the visual in Geograph (since the descriptions relate to photographs), but also implies the importance of quietness for historical interpretations of the Lake District; as the author Frederick Amadeus Malleson put it, 'external nature, with all her charms, can only occupy the mind in its leisure hours of quiet peace and meditation'. In other words, writers like Malleson discovered that they could only connect with the Lake District's beauties when they could enjoy the scenery in tranquillity.

The texts from the CLDW also contain a much more varied vocabulary than the Geograph source; more than this, comparing the use of parts of speech in the sub-corpora with the main corpora reveals that the use of all parts of speech, both in quantity and range, is statistically significantly higher in the CLDW (randomisation test, $p < 0.005$). While we did find similar numbers of relevant sentences in both corpora, with similarly rich vocabularies in terms of the number of nouns and adjectives used, the contemporary descriptions nevertheless use significantly fewer unique nouns (815 vs. 2434) and adjectives (291 vs. 922) than the historical sub-corpus in their descriptions of silence. Meanwhile, there are fewer and less varied nouns in the Geograph sub-corpus, and the quantity—though not the range—of adjectives is significantly more in our selection than in the corpus as a whole (randomisation test, $p < 0.005$). In part, the CLDW's greater linguistic variation may be attributable to the fact that the texts from this corpus tend to have longer sentences, but this difference alone seems unlikely to account for the magnitude of the change. Paying closer attention to specific types of description, as well as to individual accounts, may reveal why these changes occurred, and indicate shifts in the perception of Lakeland sounds and silences.

4.3. *Unpacking the experience of silence*

Having established that our search terms retrieved rich descriptions, we started to explore what these accounts revealed about changes to experiences of the Lakeland soundscape from the eighteenth century to the present day. We began by looking at the relationship between our search terms and the sub-corpora. Of the 66 seed terms with

Table 1. Comparison of the properties of the corpora and sub-corpora.

Corpus	CLDW		Geograph	
Version	Full corpus	Extracted silence	Full corpus	Extracted silence
Number of unique authors	55	49	1076	118
Number of sentences	43,269	590	98,206	362
Mean length of sentences (with/without punctuation)	30/26	47/41	14/13	18/16
Number of nouns (total/unique)	306,722/25,057	6271/2434	417,455/21,956	1730/815
Number of adjectives (total/unique)	102,530/9759	2595/922	122,670/10,517	848/291

which we searched, 28 returned sentences in the CLDW and 14 in Geograph. Furthermore, in Geograph only 6 terms occur more than 5 times (*quiet, peaceful, calm, peace, tranquil* and *quietly*), while in the CLDW 17 terms had a frequency greater than 5. In both corpora, *quiet* was the most prevalent term; it featured in 60% of extracted sentences in Geograph, and in 16% of sentences in the CLDW sub-corpus. *Silence* presents a more complex example. It is relatively common in the CLDW and was the 3rd most frequent search term, returned in 11% of sentences. However, in the Geograph sub-corpus, it was very rare (we found only a single occurrence). Why these differences exist requires further unpacking.

To develop a more nuanced understanding of changes in the interpretation of these terms, we began by exploring co-occurrences between our seed terms and other terms found in the extracted sentences. We firstly calculated all co-occurrences at the sentence level after removing stop words. 1904 terms co-occurred more than twice in the CLDW, while only 407 did so in Geograph, reflecting the different distributions of unique parts of speech found in Table 1. To tease out the semantics at a macro level in our texts, we identified four commonly occurring classes in the top 100 co-occurrences: references to **natural** or **anthropogenic** objects (e.g. *lake, mountain, house, road*), references to **time** (e.g. *morning, instant, time, afternoon*) and references to **generic locations** (e.g. *scene, view, spot*). In Geograph, 63 of the top 100 co-occurrences could be allocated to these classes; in the CLDW, the same was true of 51 of the top 100. We used these classes, the most common co-occurences of which are shown in Table 2, as the basis for the annotation we describe below. Many of the remaining terms described properties of objects (e.g. *old, green, little, deep*), emotions (e.g. *happy, love*) or spatial prepositions (e.g. *near, distant, close*).

Natural features were the most common class in both sub-corpora, although this is to be expected since some of these terms (e.g. *dales, tarn* and *fells* in Geograph and *lake* and *sky* in the CLDW) occur equally or more often in the corpora as a whole. Anthropogenic terms associated with sound are dominated in Geograph by *road*, and in particular the *M6*, a *motorway* which runs along the edge of the National Park. In the CLDW the significant co-occurrences also relate to transport, but suggest movement by

Table 2. Classified co-occurrences. Words denoted with an asterisk occur significantly more often in the silence sub-corpora than in the random subsets of the corpora (randomisation test, $p < 0.005$). Multiple values (e.g. 11 +10 + 9) denote co-occurrences with more than one seed-term.

	CLDW		Geograph		CLDW		Geograph	
Corpus	word	count	word	count	word	count	word	count
Class	Natural				Anthropogenic			
	nature*	11	lake*	15	house	8	road*	24
	lake	11 + 10 + 9	valley*	15	man*	7	park	17
	vale*	9	dales*	9	boat*	7	lane	10
	clouds*	8	tarn	8	gentlemen	6	M6*	10
	sky	7	fells	8	town	6	motorway*	9
Class	Generic locations				Time			
	way	9	spot*	14	day*	12 + 8	day*	12 + 10
	country	8	place*	14	evening*	7	morning*	10
	scene*	8 + 7	district*	13	time	6 + 6	early*	8
	spot	7	corner*	12	instant*	6	times*	8 + 8
	paradise	7	area	10	morning*	6	Sunday*	7

boat, and also indicate human habitation. By contrast, generic place descriptions (e.g. *spot, place* and *scene*) are often used to characterise locations which are discussed with respect to sound. These descriptions seem to imply not only aural, but visually perceptible locations (e.g. a quiet spot or a peaceful valley). In Table 2 we do not, for reasons of space, show the seed terms with which co-occurrences occurred. For Geograph, these are dominated by *quiet*. Other terms, such as *calm* and *peace*, are also prominent in the CLDW. *Calm* in particular is often found in conjunction with weather and water-related terms. As we will see in more detail shortly, calmness implies not only an absence of noise, but also movement.

We have reached, for now, the limits of what can be achieved by slicing and dicing our corpora and, guided by these observations, now resort to alternative methods to analyse the nature of the references to silence and tranquillity in the CLDW and Geograph. As we will now explain, in order to annotate the two sub-corpora, we combined the results from this macroanalysis with a micro-analytic approach adopted from literary studies. In this way, our annotation agreement is perhaps the most cohesive evidence of a middle ground practice which establishes a dataset that is especially suited for the kind of multiscalar, multidisciplinary textual analysis for which we advocate here.

4.4. *Characterising the silence*

To better understand the nature of our two corpora, we developed a two-layer annotation scheme. The first layer of annotation aimed to capture the nature of the sounds and sound-related descriptions found in our texts. Based on our macroreading, we proposed the following broad classifications:

- **Total silence and tranquil sounds**: Either explicit descriptions of complete silence or a combination of tranquil sounds without contrast (e.g. *the silence was total*).
- **Contrasting sounds**: Descriptions capturing ephemerality in silence or tranquillity at a location, comparing one location to another that is less tranquil, or mentioning a sound which adds (or detracts) from the overall tranquillity (e.g. *we heard nothing but the hum of the bees*).
- **Combination of visual and aural**: Silence is implicit in the overall description of a scene, and visual properties are also conveyed (e.g. *a quiet spot above the lake*).
- **No movement**: Implied silence or tranquillity, but explicit mention of a lack of movement (e.g. *yachts sit at anchor in this quiet bay*).
- **Not relevant**: Search terms used in another sense, sounds which do not convey silence or tranquillity or descriptions of sounds not situated in the landscape (e.g. *the clock ticked loudly*).

The second layer of our annotation scheme relates to the nature of the potential sound emitters in a description. Here, we follow Krause (2008) in that, where an explicit mention of a potential sound emitter was made, we allocated it to one, or a combination thereof, of the following classes:

- **Geophony**: Natural sounds produced by non-biological sources (e.g. *wind, thunder, waterfalls*).

- **Biophony**: Natural sounds produced by animals (e.g. *lowing cows* or *humming bees*).
- **Anthrophony**: Sounds produced by humans either directly or indirectly (e.g. *noisy kids* or *busy road*).

Annotation of text is often challenging, and the texts in the CLDW were particularly difficult to interpret. To mitigate the texts' ambiguities as much as possible, we carried out an iterative annotation process. Two of the authors annotated 10% of each sub-corpora, discussing disagreements and refining unclear guidelines. After three iterations (i.e. annotating 30% of both sub-corpora), an inter-annotator agreement (Cohens Kappa) of 0.88 for types of silence and 0.90 for sound emitters in the Geograph corpus was reached. According to Landis and Koch (1977) this level of agreement is 'almost perfect', and a single annotator then annotated the remaining 70% of Geograph texts. For the CLDW, after three rounds of iteration we plateaued at 'substantial' inter-annotator agreement of 0.62 (type of silence) and 0.6 (sound emitter) respectively. Both annotators therefore annotated the remaining 70% of CLDW texts, and for cases where annotations differed discussed the texts until we reached a consensus. Table 3 shows the first layer of our annotation as absolute counts.

Three main characteristics are striking when the texts are processed in this way. First is the almost complete absence of total silence and tranquil sounds in Geograph, suggesting—as we observed above—that the lack of descriptions using the search term *silence* was indeed indicative of a change in the way the Lake District soundscape is perceived. Second is the much larger proportion of extracted descriptions found in the CLDW which were not relevant for an inquiry into acoustic experiences. This is despite our use of a historical thesaurus, which we expected to be more effective at extracting descriptions from the CLDW than Geograph. There are, we think, two reasons for this result. Firstly, Geograph descriptions are less complex and more literal. Secondly, there is a demonstrably diachronic variation in language, illustrated by the ambiguity of our search terms with respect to the CLDW (e.g. *quiet*: 37%, *peace*: 69%, *quietly*: 72%). This variation once more highlights the importance of our interdisciplinary approach, and the importance of a microreading of the texts.

The third key characteristic of this first layer of annotation concerns the overall distribution of classified sounds. The overall ranking, if not the proportion, is the same for both sub-corpora; descriptions of a combination of visual and aural are most common, followed in decreasing rank by contrasting sounds, no movement, total silence and tranquil sounds.

Figure 2 illustrates the distribution of potential sound emitters as a function of our sound classes. Biophony is relatively uncommon in both sub-corpora, whereas the presence of geophony in all the classes demonstrates the importance of the physical

Table 3. Counts of descriptions per class.

Corpus/Class	CLDW	Geograph
Total silence and tranquil sounds	46	3
Contrasting sounds	70	108
Combination of visual and aural	168	179
No movement	48	40
Not relevant	258	32

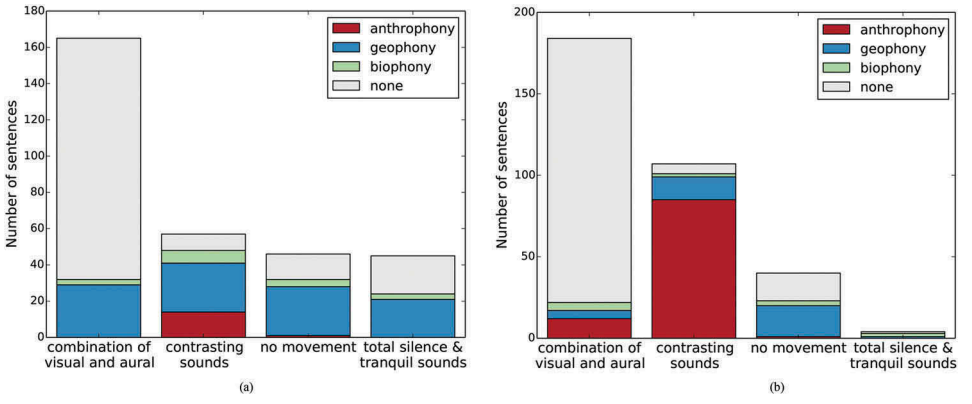

Figure 2. Number of sentences per silence class grouped by sound emitters: (a) CLDW (b) Geograph.

landscape and weather as perceived emitters of tranquil sounds in the Lake District. Yet the almost complete absence of anthrophony in the CLDW, and its dominance in the Geograph 'contrast class', stands out. This difference, we suggest, can be traced back to William Wordsworth, who constructed a certain type of cultural soundscape that affirmed the Lake District's status as a region of peace and tranquillity which was increasingly unusual in industrialising and urbanising Britain (Taylor 2018). Wordsworth deliberately overlooked noises that contrasted with the tranquil cultural soundscape he preferred; the constant cacophony from the mining industry that proliferated in the Western Lakes, like the quarries and other sites of industrial development that could be found throughout the region, are 'artfully' ignored in his works (Levinson 1986). Instead, Wordsworth's writing encouraged generations of writing about and visitors to the Lake District to focus on the sounds of flowing water and breezes blowing through the leaves on trees (our geophony), rather than the more realistic—but less pastoral—sounds of industrial activity. In Geograph, by contrast, tranquillity is valued not by ignoring the intrusion of anthropogenic influences, but rather explicitly contrasting tranquil locations to anthrophonic intrusions.

The importance of geophony, rather than anthrophony, in the CLDW is emphasised by the word clouds we created to show more clearly the patterns in the class of contrasting sounds in Figure 3. The word clouds showed the 150 most frequent nouns in the two sub-corpora, and the importance of geophony in the CLDW is underlined through the prominence of terms such as *thunder, cataracts* and *waterfalls*. A previously documented interest in echoes in this period is also visible

Figure 3. 150 most frequent nouns extracted from the class 'contrast': (a) CLDW (b) Geograph.

(Taylor 2018). In Geograph, the prominence of transport as a potential sound emitter is again clear (e.g. *road, M6, motorway* and *car*), while there is also evidence of the presence of other people as a source of discordant sounds (e.g. *visitors* and *walkers*).

The most significant finding from our macroanalysis and the resultant annotation is the revelation that what is meant by quietness has undergone a significant shift in the intervening years between the two corpora. How that features in our texts, and why that might be the case, requires more detailed focus on key individual texts within the wider collective on which we have focused so far.

4.5. *Reading the silence*

Once we had established these annotation rules, we applied them in order to ascertain quantitatively whether these terms possessed positive or negative connotations in our sub-corpora. We assigned mean sentiment values to each description using an existing Opinion Lexicon (Hu and Liu 2004) and a pretrained set of GloVe word embeddings (Pennington *et al*. 2014) to attach sentiment values to words not contained in the lexicon. Figure 4 shows histograms of sentiment for randomly selected sentences from both corpora, as well as our two sub-corpora. As the bias towards the right in the histograms indicates, descriptions of quietness tend to be positive. Secondly, both of our sub-corpora are statistically significantly more positive (t-test, $p < 0.005$) than the corpora from which they are extracted. This difference is much more marked for Geograph (overall corpus mean sentiment 0.80 ± 0.63 vs. silence sub-corpus mean sentiment 1.90 ± 0.59) than in the CLDW (overall corpus mean sentiment 1.58 ± 0.59 vs. silence sub-corpus mean sentiment 1.80 ± 0.49).

To test if this bias towards positive values can be explained by the presence of our seed words in the descriptions of silence, we calculated the sentiment values for descriptions without taking into account our seed words. The absolute mean value of the difference between the original sentiments and sentiments without seed words is small (0.4 in Geograph and 0.2 in the CLDW). Therefore, we concluded that such descriptions are in general associated with positive sentiment.

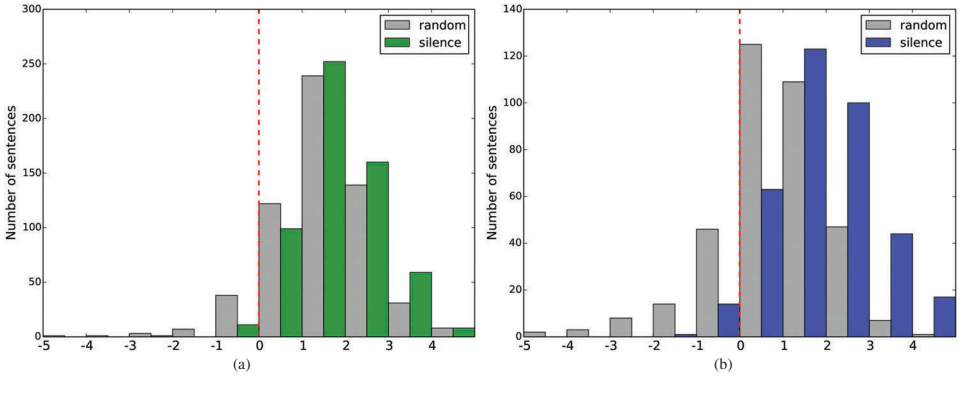

Figure 4. Number of sentences grouped by sentiment values and compared against random sample (a) CLDW $n_{sentences} = 590$, t-test $p < 0.005$, (b) Geograph $n_{sentences} = 362$, t-test $p < 0.005$.

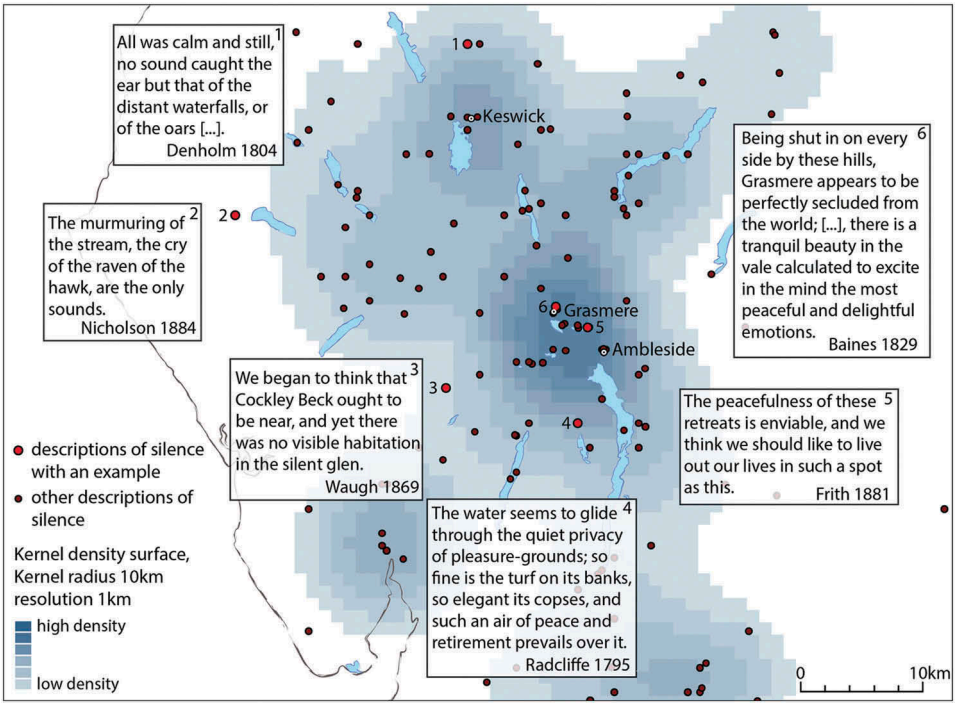

Figure 5. Locations of texts in the Lake District associated with toponyms in the CLDW and an associated kernel density surface.

The advantages—as well as the limits—of our approach in general are perhaps best illustrated by this sentiment analysis. We find that our text snippets are generally positive, and descriptions related to silence in Geograph are more positive than the corpus as a whole. The assumption—made by us and by the Opinion Lexicon—that quietness and peace relate to tranquillity, and that these are desirable qualities in a rural landscape, seems common-sensical today (MacFarlane *et al*. 2004). Evaluating why this is the case, though, relies on closer attention to individual accounts within the corpora.

We saw earlier that it has certainly not always been the case that quietness was a desirable feature (Fisher 1999), and the greater linguistic variety of the CLDW indicates that quietness was both a more common and a more complex phenomenon. William Gilpin, the Cumbrian curate most famous for his development of the picturesque mode of landscape evaluation, is influential over the promotion of quietness in the Lake District (Taylor 2018). In his *Observations, Relative Chiefly to Picturesque Beauty* (1786), Gilpin wrote of Lorton vale in the north-western Lakes that it was a place that could 'pretend not to dignity'; it could only aspire to be a 'mere [scene] of tranquillity'. Nevertheless, such a place held its own charms for Gilpin, not least because they had the potential to transport him into a particular mental state. He continued that he 'might have wished for a quiet, tranquil hour, when the glimmering surfaces of things are sometimes perhaps more pleasing – at all times certainly more soothing, than images of the brightest hue'. Gilpin grammatically links *quiet* with *tranquil* here. Indeed, the lack of sound and movement in

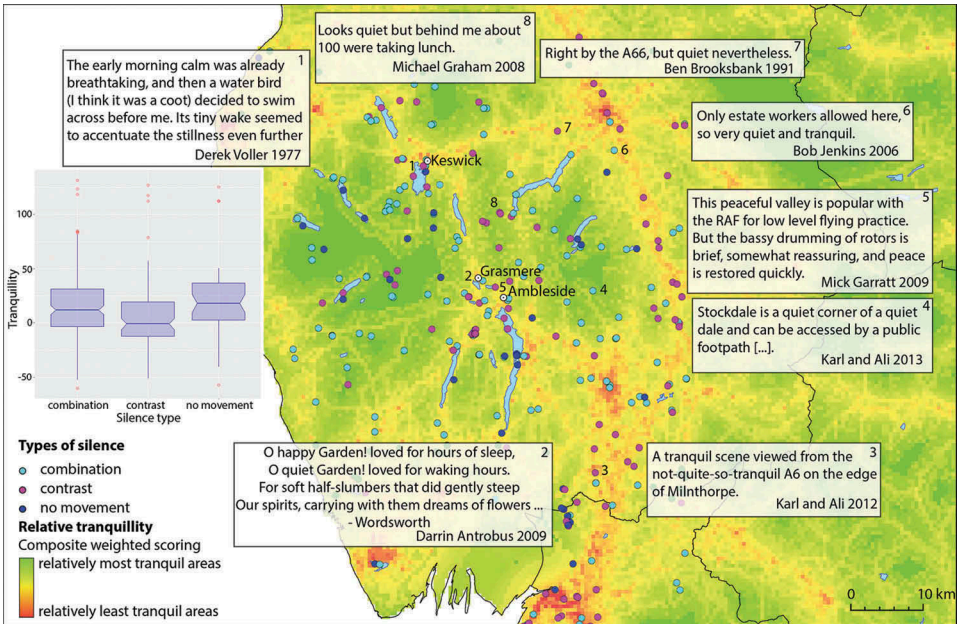

Figure 6. Comparison of the map of relative tranquillity MacFarlane 2004 and types of silence extracted from the Geograph corpus. National Tranquillity mapping data 2007 developed for the Campaign to protect rural England and Natural England by Northumbria University. OS Licence number 100018881.

the surrounding area seems to calm his mind in ways that allow him to connect more effectively with the surrounding landscape; it is only in this quiet and tranquil state that Gilpin can appreciate the 'glimmering surfaces of things' around him.

This sense of quietness as indicating acoustic peace, physical stillness and—crucially—a closely related sense of mental calm was inherited by later Lakeland authors. Joseph Mawman, writing a couple of decades after Gilpin, discovered a similar 'harmony with the soothing quiet which prevailed all around'. As for Gilpin, this quietness established for Mawman a sense of what we might recognise today as mindfulness: the quiet 'disposed us,' he wrote, 'to reflect seriously upon that interminable question, 'What state of life is best fitted for happiness?' Later still, the Lancashire poet Edwin Waugh was even more explicit: he thought that '[g]oing from a crowded city into this little monastic town [Cartmel] is almost like going to bed, or sinking into an antiquarian dream, – all is so quaint and quiet'. For Waugh, quietness does not simply establish a calm and meditative state; it actually feels like going back in time to an earlier moment when the world was, he imagines, less frantic.

This meaning of *quiet* as indicating peacefulness in the landscape as well as a form of mental calm is not remarked on by the Geograph writers. That is not, however, to say that the connection does not exist. Karl and Ali, for instance, describe Stockdale as 'a quiet corner of a quiet dale', and here the repetition of the search term highlights the writers' enhanced sense of being removed from the busier, more inhabited parts of the region, and emphasises a certain stasis in the valley. Bob Jenkins, meanwhile, implies a Gilpinian quietness when he writes of Lowther Park that since 'only estate workers

[are] allowed', it is 'very quiet and tranquil' (Figure 6). While in the CLDW, the writers needed to explain this link between the physical soundscape and their own mental state, it is taken for granted in the Geograph texts. These later authors therefore require less diverse language to indicate the sense of tranquillity they discover in the Lake District's quiet places; they, like us, have inherited from writers such as those represented in the CLDW a sense that quietness is a positive value. The difference is that, for the Geograph authors, quietness is a much scarcer commodity than it was for the CLDW writers (even at a historical moment that, elsewhere, was witnessing an exponential increase in noise (Picker 2003, Bour 2016)). Nevertheless, as Timothy Morton writes, 'in order to differentiate [quietness] there must be some roughness, some *noise*' (2009, p. 71), and the writers in both our corpora share a sense of relief at discovering spaces of quiet in an increasingly noisy world.

4.6. *Mapping the silence*

In a final exploration of our corpora, we set out to discover not only what was talked about, but where. Place names in the CLDW have been georeferenced using toponym recognition and resolution (Rayson *et al*. 2017) and used to explore broader themes, including local variation in the use of aesthetic language (Donaldson *et al*. 2017) and acoustic experience (Taylor *et al*. 2018). In contrast, the Geograph corpus explicitly links descriptions to 1km grid squares, and texts can therefore be mapped without any additional processing. Since the two corpora document space in different ways, and at very different scales, the spatial extents which can be associated with texts are not directly comparable. In the following, we seek to once more find a middle ground, and map the two corpora to space in ways which are appropriate given these considerations.

The toponyms in the CLDW situate the texts in space, but it is often not possible to assign coordinates to the areas they describe with much accuracy. Associating extracted sentences and related sounds with locations requires us not only to identify corresponding toponyms, but also to specify an appropriate document scope that relates the content to a location (Andogah *et al*. 2012). However, the complexity of the sentence structures in the CLDW frequently makes this process challenging. For example, when James Denholm describes movement on Derwentwater—the lake beside the town of Keswick—the closest toponym actually refers to a nearby mountain summit: 'the hills upon the left were in the shade, as was the mountain of Skiddaw, lying, together with the islands, directly in front. All was calm and still, no sound caught the ear but that of the distant waterfalls, or of the oars, striking, in alternate succession […] the surface of the lake'. Even where we can identify a named feature, associating it with a spatial extent is still challenging. On what part of Derwentwater did Denholm row, for instance, and at what point on his journey did he reflect on that state of calmness? Nevertheless, these toponyms do allow us to identify the area being discussed, even if we often cannot map it with any great degree of specificity.

In light of this complexity, we used a simple heuristic to associate toponyms with sentences. We first identified any toponyms in the target sentence; if none were found, we looked firstly one sentence back, and then one sentence forward in order to identify the most appropriate document scope. Where more than one toponym was found, the sentence was associated with multiple locations. For the 332 descriptions of silence

extracted from the CLDW, we could map a little less than half (157) to locations in this way. Given the above limitations, we used a kernel density estimation with a coarse kernel of 10km that reflects the uncertainty registered in the texts, but still allows us to explore spatial patterns in the data (Figure 5).

We can see from Figure 5 that Grasmere emerges in this kind of analysis as being particularly significant for discussions of sound in the CLDW. Grasmere's importance is, in some ways, not surprising: it was located on the main route between Ambleside and Keswick, and as a result was an almost inevitable sight on any Lakeland tour in the period (Murrieta-Flores *et al*. 2017). Yet, the texts in the CLDW point to another reason why Grasmere was considered to be a particularly evocative location for the pursuit of quiet repose. The poet William Wordsworth (1770–1850) resided in or near the village from 1799 until his death, and he celebrated the peace he enjoyed there in his writing. More than that, though, he also highlighted certain acoustic experiences: the echo of his sister-in-law's laugh around the valley, for instance, encouraged several imitators (Taylor 2018). Loughrigg Tarn also became a popular excursion for tourists in pursuit of the sounds of Wordsworthshire (Donaldson *et al*. 2015).

Wordsworth's influence can, in fact, be traced through to the modern day: one of the contemporary descriptions of Grasmere quotes the poet verbatim (Figure 6) to evoke the peace that allowed Wordsworth—and, perhaps, his modern reader—to enjoy 'soft half-slumbers' in the tranquil valley. The greater degree of precision offered by, and the contemporary relevance of, the Geograph corpus meant that we were able to compare this data directly with MacFarlane's *et al*.'s (2004) tranquillity study. To do so, we resampled the 500m *map of relative tranquillity* to the 1km resolution of Geograph using bilinear interpolation. Figure 6 shows the locations of the three most prominent types of silence found in these data (c.f. Table 3) along with box plots of tranquillity values. From this, it seems that contrasting sounds are associated with low values of tranquillity, particularly near the M6. The emergence of this area indicates that places associated with silence and tranquillity have 'spread' from the central Lake District that had been the focus for nineteenth-century travellers. Instead, today's visitors find that almost the entire National Park offers a sense of tranquil quiet.

A similar outcome occurs when we analyse this corpus quantitatively. Using a non-parametric Kruskal-Wallis test, we analysed tranquillity as a function of silence. A significance level of $p < 0.01$ was used to reject the null hypothesis that all types of silence were associated with similarly distributed values of tranquillity. Since this test was significant, we used a post-hoc Dunn test to compare all tranquillity values with a significance level of $p < 0.01$[6] (Table 4) (Dunn 1961).

These tests revealed that contrasting sounds in the Geograph texts are statistically significantly associated with lower values of relative tranquillity than both combination and no movement, based on an independently created model (MacFarlane *et al*. 2004). This link suggests that, firstly, the prominence of anthrophony in contrasting sounds (Figure 2) reflects real variation in environmental properties, since low values of tranquillity are typically associated with anthropogenic disturbance. Secondly, the comparison demonstrates that the texts we extracted are in broad agreement with an independently produced model. Thirdly, it also demonstrates a significant strength of our multidisciplinary approach: this method has allowed us to identify locations which

Table 4. Adjusted significance values for a post-hoc Dunn test comparing distributions of tranquillity values associated with types of silence.

Comparison	*p*.adj
Contrasting sounds vs. Combination of visual and aural	0.003
No movement vs. Combination of visual and aural	0.247
Contrasting sounds vs. No movement	0.002

are considered tranquil despite an unpromising setting. For example, Ben Brooksbank's describes a scene '[r]ight by the A66' that is 'quiet nevertheless.' Brooksbank is indicative of a significant group of authors in the Geograph corpus who identify locations beside busy roads as being comparatively peaceful. It seems that, for the modern visitor, complete quiet is not necessary for the discovery of tranquillity.

5. Discussion

Our aim in this project was to extract and interpret descriptions of, and diachronic and spatial variation in, perceived silence from historical and contemporary textual descriptions. By adopting a blend of methods, focussing on detailed reading of individual texts, annotation and stratification of descriptions of silence and a range of quantitative analyses of both corpora we were able to fulfil this aim.

Although both corpora contained references to our silence-related seed words, these were much more prominently used by authors in the CLDW (89%) than Geograph (10%). This, we argue, reflects the importance of peace and silence in historical accounts of the Lake District. Descriptions of silence were also, in the historic texts, associated with a richer use of both nouns and adjectives. This reflects, on the one hand, the more literal nature of the short descriptions in Geograph and, on the other, the need to set out the authors' mental state in a description of silence or peace in the CLDW.

To find silence-related descriptions we used a variety of seed terms. These also demonstrated a clear diachronic change, with descriptions of total silence or calmness almost totally disappearing in Geograph. Exploring terms which co-occurred with our seed terms helped us identify changes in the nature of terms associated with silence. Here we see two changes over time. Firstly, nouns associated with transport (e.g. *road, motorway*) emerge as common co-occurrences in Geograph descriptions of silence. Secondly, we note that generic place descriptions (e.g. *spot, place, corner*) become increasingly important, reflecting perhaps a change from the description of a whole landscape, to a specific location within it.

These first explorations of our corpora guided the following classifications of both silence and related sound emitters. Having identified four key classes of sounds, we annotated the extracted descriptions. This annotation further demonstrated the almost complete absence of total silence and tranquil sounds in the contemporary data, and also showed the increased importance of silence expressed through contrast. By annotating sound emitters, we identified the concern about anthropogenic disturbances in the modern landscape. Both corpora privilege descriptions of geophony over biophony, and in doing so adhere to a version of the cultural soundscape that can be traced back to writers like Wordsworth.

Wordsworth's influence on historical descriptions of the Lake District is clearly visible in the general positivity associated with descriptions of this landscape and silence within

it (Figure 4). By contrast, Geograph descriptions are in general neutral, reflecting the aim of the collection to describe the landscape. Nonetheless, descriptions referring to silence are statistically significantly more positive, reflecting value given to silence as a cultural resource. By projecting our descriptions into space, the persistent influence of Wordsworth is emphasised. In both corpora, we find a cluster of descriptions centred around Grasmere, a location popularised and written about by Wordsworth and his followers. Comparing contemporary descriptions of silence to a map of relative tranquillity showed that contrast is both semantically and spatially associated with anthropogenic disturbance. This comparison also illustrates how our textual descriptions can indeed allow us to identify tranquil locations even in busy areas of the landscape.

Our aim in this work was to uncover a middle ground that combines interdisciplinary methods to generate multiscalar perspectives on textual, spatial data. Pragmatically, if we wish to make a contribution to Landscape Character Assessment, this result matters since it demonstrates two key points. Firstly, the prominence of descriptions in our contemporary corpus which refer to generic places (e.g. *spot, place, corner*) implies a form of landscape perception that focusses on locations with some form of gestalt coherence (Schroeder 2007). Secondly, modelling relative tranquillity is contrary to current GIS-based attempts at quantifying such properties, which often focus on distance from potential emitters as a proxy for disturbance (e.g. Carver *et al*. 2002, MacFarlane *et al*. 2004). Rather, our approach suggests an additional need for modelling tranquil places by contrast, as suggested—though in a very different context—by Winter and Freksa (2012). Further, this approach points to an oft-observed dichotomy between attempts to model landscape properties as continuous fields (Mücher *et al*. 2010) and the diverse ways in which people perceive and categorise the world (Mark *et al*. 2011).

It is, of course, important to note a number of limitations with our approach. Firstly, our results are dependent on the choices we made during preprocessing, including: the seed words selected; the reliability of our annotation; and the specific methods we used (e.g. the quality of the part of speech tagging, the use of GloVe embeddings and our approach to sentiment analysis). However, though such limitations are part and parcel of any text-based approach, we argue that our results are robust since quantitative macroreadings of our corpora were interpreted through, and substantiated by, qualitative microreadings. Secondly, our corpora have different properties, particularly with respect to georeferencing and granularity. Putting aside the inevitable uncertainty introduced by mapping toponyms directly to point locations, the rich descriptions found in the CLDW cannot be easily mapped to areas associated with places described in the texts (Murrieta-Flores *et al*. 2017). We suggest that until methods such as those proposed by Moncla *et al*. (2016) can be applied successfully to historical texts, spatial comparisons of this kind are best performed on the region as a whole (c.f. Figure 2).

6. Conclusions

We set out to explore how finding the middle ground—a place for a blend of methods from a range of disciplines—could offer us insights into two temporally distinct, spatially overlapping corpora describing experiences of the Lake District landscape. In particular:

- Unstructured texts offer rich, semantically diverse, and spatially groundable insights into landscape perception, and more generally access to understandings of the way place is made and conceptualised.
- Diachronic use of corpora offer insights into ways in which readings of contemporary and historical landscape descriptions are intertwined.
- Spatial contiguous models of properties such as tranquillity can be enhanced and refined through complementary analysis of spatially grounded textual sources.

Our results do not necessarily suggest new ways of understanding silence and soundscapes. Rather, they reveal scalable approaches towards exploring how people represent, in writing, their individual experiences of landscapes in given places. Practically speaking, our approach suggests ways of extracting and analysing important information required in Landscape Character Assessment, and could be scaled up to cover large spatial extents. More generally, we suggest that GIScience would do well to consider the opportunities offered by critically exploring rich unstructured text, whilst literary historical studies should embrace the plethora of authors and viewpoints offered by this kind of approach. For both disciplines, this middle ground offers a way of increasing the breadth of participation in the production of spatial information and knowledge.

Notes

1. https://www.nps.gov/subjects/sound/soundsmatter.htm.
2. https://github.com/UCREL/LakeDistrictCorpus.
3. http://www.geograph.org.uk/.
4. Based on an anonymous survey carried out by the initiators of the project.
5. https://www.nltk.org/.
6. p value adjusted for multiple means using the Benjamini-Yekuteili method https://www.rdocumentation.org/packages/FSA/versions/0.8.20/topics/dunnTest.

Acknowledgments

We would like to gratefully acknowledge all the contributors to Geograph British Isles (Creative Commons Attribution-ShareAlike 2.5 License). We are grateful to Graeme Willis (Campaign to Protect Rural England) and Nick Groome (Ordnance Survey) for their help in accessing the National Tranquillity Mapping Data. Finally, we thank the reviewers for their useful suggestions.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Research for this article was partly supported by the Leverhulme Trust-funded project 'Geospatial Innovation in the Digital Humanities: A Deep Map of the English Lake District' (RPG-2015-230). IGs contribution to this paper received funding from the European Research Council (ERC) under the European Unions Seventh Framework Programme (FP7/2007-2013)/ERC grant "Spatial Humanities: Texts, GIS, places" (agreement number 283850). RSP gratefully acknowledges support from the

Swiss National Science Foundation Project EVA (166788) and the University Research Priority Programme on Language and Space.

Notes on contributors

Olga Chesnokova is a PhD student at the University of Zurich. Her focus lies in the use of unstructured text to better understand landscape and landscape perception.

Joanna E. Taylor is Presidential Academic Fellow in Digital Humanities at the University of Manchester. Her research uses digital humanities methodologies – particularly digital mapping – to explore literary geographies and environmental histories from the long nineteenth-century.

Ian N. Gregory is Professor of Digital Humanities at Lancaster University. His research interests lie in the Spaital Humanities, and particularly how GISc techniques can be used in combination with other digital humanities approaches, to analyse large bodies of text.

Ross S. Purves is a Professor at the Department of Geography of the University of Zurich. His research interests focus on how we can answer and explore societally relevant geographic questions paying attention to vagueness and uncertainty, often using unstructured data in the form of text as a primary source.

ORCID

Olga Chesnokova http://orcid.org/0000-0002-4298-1789
Joanna E. Taylor http://orcid.org/0000-0001-8597-0097

References

Agnew, V., 2012. Hearing things: music and sounds the traveller heard and didn't hear on the grand tour. *Cultural Studies Review*, 18 (3), 67–84. doi:10.5130/csr.v18i3.2855

Ammer, C., 2013. *The American heritage dictionary of idioms*. 2nd ed. Boston, MA: Houghton Mifflin.

Andogah, G., Bouma, G., and Nerbonne, J., 2012. Every document has a geographical scope. *Data & Knowledge Engineering*, 81–82, 1–20. doi:10.1016/j.datak.2012.07.002

Antoniou, V., Morley, J., and Haklay, M., 2010. Web 2.0 geotagged photos: assessing the spatial dimensions of the phenomenon. *Geomatica*, 64 (1), 99–110.

Bour, I., 2016. Foreword: noise and sound in the Eighteenth century. *In*: *Etudes Epistémè: revue de littérature et de civilisation (XVIe-XVIIIe siecles)*, 29. Available from: http://journals.openedition.org/episteme/11 [Accessed 15 April 2018]. doi:10.4000/episteme.1136

Burenhult, N. and Levinson, S.C., 2008. Language and landscape: a cross-linguistic perspective. *Language Sciences*, 30 (2–3), 135–150. doi:10.1016/j.langsci.2006.12.028

Butler, A., 2016. Dynamics of integrating landscape values in landscape character assessment: the hidden dominance of the objective outsider. *Landscape Research*, 41 (2), 239–252. doi:10.1080/01426397.2015.1135315

Butler, J.O., *et al*., 2017. Alts, Abbreviations, and AKAs: historical onomastic variation and automated named entity recognition. *Journal of Maps and Geography Libraries*, 13, 58–81. doi:10.1080/15420353.2017.1307304

Carver, S., Evans, A., and Fritz, S., 2002. Wilderness attribute mapping in the United Kingdom. *International Journal of Wilderness*, 8 (1), 24–29.

Chesnokova, O. and Purves, R.S., 2018. From image descriptions to perceived sounds and sources in landscape: analyzing aural experience through text. *Applied Geography*, 93, 103–111. doi:10.1016/j.apgeog.2018.02.014

Coates, P.A., 2005. The strange stillness of the past: toward an environmental history of sound and noise. *Environmental History*, 10 (4), 636–665. doi:10.1093/envhis/10.4.636

Comber, A., Fisher, P., and Wadsworth, R., 2005. What is land cover? *Environment and Planning B: Planning and Design*, 32 (2), 199–209. doi:10.1068/b31135

Corbin, A., 1986. *The foul and the fragrant: odor and the French social imagination*. USA: Harvard University Press.

Council of Europe, 2000. European landscape convention. *Report and Convention Florence*, ETS No. 17 (176), 8.

Donaldson, C., Gregory, I.N., and Murrieta-Flores, P., 2015. Mapping 'Wordsworthshire': a GIS study of literary tourism in Victorian Lakeland. *Journal of Victorian Culture*, 20 (3), 287–307. doi:10.1080/13555502.2015.1058089

Donaldson, C., Gregory, I.N., and Taylor, J.E., 2017. Locating the beautiful, picturesque, sublime and majestic: spatially analysing the application of aesthetic terminology in descriptions of the English Lake District. *Journal of Historical Geography*, 56, 43–60. doi:10.1016/j.jhg.2017.01.006

Dubois, P., 2016. The impossible temptation of noise in late Eighteenth-century English music. *In*: *Etudes Epistémè: revue de littérature et de civilisation (XVIe-XVIIIe siecles)*, 29. Available from: https://journals.openedition.org/episteme/1122 [Accessed 15 April 2018]. doi:10.4000/episteme.1122

Dunn, O.J., 1961. Multiple comparisons among means. *Journal of the American Statistical Association*, 56 (293), 52–64. doi:10.1080/01621459.1961.10482090

Fairclough, G. and Herring, P., 2016. Lens, mirror, window: interactions between historic landscape characterisation and landscape character assessment. *Landscape Research*, 41 (2), 186–198. doi:10.1080/01426397.2015.1135318

Fisher, J.A., 1999. The value of natural sounds. *Journal of Aesthetic Education*, 33 (3), 26–42. doi:10.2307/3333700

Fisher, P., Wood, J., and Cheng, T., 2004. Where is Helvellyn? Fuzziness of multi-scale landscape morphometry. *Transactions of the Institute of British Geographers*, 29 (1), 106–128. doi:10.1111/tran.2004.29.issue-1

Fuchs, R., *et al*., 2015. The potential of old maps and encyclopaedias for reconstructing historic European land cover/use change. *Applied Geography*, 59, 43–55. doi:10.1016/j.apgeog.2015.02.013

Gregory, I. and Donaldson, C., 2016. Geographical text analysis: digital cartographies of Lake District literature. *In*: D. Cooper, C. Donaldson and P. Murrieta-Flores, eds. *Literary mapping in the digital age*. London: Routledge, 67–78. ISBN 9781472441300.

Grenon, P. and Smith, B., 2004. SNAP and SPAN: towards dynamic spatial ontology. *Spatial Cognition and Computation*, 4 (1), 69–104. doi:10.1207/s15427633scc0401_5

Hegarty, P., 2007. *Noise/Music: a history*. London: Continuum International Publishing Group.

Hu, M. and Liu, B., 2004. Mining and summarizing customer reviews. *Proceedings of the 2004 ACM SIGKDD international conference on knowledge discovery and data mining KDD 04*, Seattle, WA, USA, 04, 168.

Jockers, M., 2013. *Macroanalysis: digital methods and literary history*. Chicago: University of Illinois Press.

Joy, L., 2014. Relative obscurity: the emotions of words, paint and sound in Eighteenth-century literary criticism. *History of European Ideas*, 40 (5), 644–661. doi:10.1080/01916599.2013.860291

Krause, B., 2008. Anatomy of the Soundscape. *Journal of the Audio Engineering Society*, 56, 1/2.

Lake, M.W., Woodman, P.E., and Mithen, S.J., 1998. Tailoring GIS software for archaeological applications: an example concerning viewshed analysis. *Journal of Archaeological Science*, 25 (1), 27–38. doi:10.1006/jasc.1997.0197

Landis, J.R. and Koch, G.G., 1977. The measurement of observer agreement for categorical data. *Biometrics*, 33 (1), 159–174. doi:10.2307/2529310

Levinson, M., 1986. *Wordsworth's great period poems: four essays*. Cambridge: Cambridge University Press.

Leyk, S., Boesch, R., and Weibel, R., 2006. Saliency and semantic processing: extracting forest cover from historical topographic maps. *Pattern Recognition*, 39 (5), 953–968. doi:10.1016/j.patcog.2005.10.018

Lowenthal, D., 1976. Tuning into the past: can we recapture the soundscapes of bygone days?. *UNESCO Courier*, November, (29), 15–20.

MacFarlane, R., *et al*., 2004. *Tranquillity mapping: developing a robust methodology for planning support*. Technical report on Research in the Northumberland National Park and the West Durham Coalfield, Northumbria University, Newcastle University.

Manning, C.D. and Schutze, H., 1999. *Foundations of statistical natural language processing*. Cambridge, MA: The MIT Press.

Mark, D.M., *et al*., eds., 2011. *Landscape in language. Transdisciplinary perspectives*. Amsterdam/ Philadelphia: John Benjamins.

McGarigal, K. and Marks, B.J., 1994. FRAGSTATS: spatial pattern analysis program for quantifying landscapesStructure. *General Technical Report PNW-GTR-351. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station*. Portland, OR, 97331 (503), 134. doi:10.3168/jds.S0022-0302(94)77044-2

Miller, N.P., 2008. US National parks and management of park soundscapes: a review. *Applied Acoustics*, 69 (2), 77–92. doi:10.1016/j.apacoust.2007.04.008

Moncla, L., *et al*., 2016. Reconstruction of itineraries from annotated text with an informed spanning tree algorithm. *International Journal of Geographical Information Science*, 30 (6), 1137–1160. doi:10.1080/13658816.2015.1108422

Morton, T., 2009. *Ecology without nature: rethinking environmental aesthetics*. Harvard: Harvard University Press.

Mücher, C.A., *et al*., 2010. A new European Landscape Classification (LANMAP): a transparent, flexible and user-oriented methodology to distinguish landscapes. *Ecological Indicators*, 10 (1), 87–103. doi:10.1016/j.ecolind.2009.03.018

Murrieta-Flores, P., Donaldson, C., and Gregory, I.N., 2017. GIS and literary history: advancing digital humanities research through the spatial analysis of historical travel writing and topographical literature. *Digital Humanities Quarterly*, 11, 1.

Nomination, 2015. Nomination of the English Lake District for inscription on the world heritage list. [online]. Available from: https://whc.unesco.org/en/list/422 [April, 7, 2018].

Pennington, J., Socher, R., and Manning, C.D., 2014. GloVe: global Vectors for Word Representation. *In*: *Conference on empirical methods in natural language processing (EMNLP 2014),* Doha, Qatar.

Picker, J.M., 2003. *Victorian soundscapes*. Oxford: Oxford University Press.

Pickles, J., ed., 1995. *Ground truth: the social implications of geographic information systems*. New York: Guilford Press.

Pijanowski, B.C., *et al*., 2011. Soundscape ecology: the science of sound in the landscape. *Bioscience*, 61 (3), 203–216. doi:10.1525/bio.2011.61.3.6

Quercia, D. and Schifanella, R., 2015. Smelly maps: the digital life of urban smellscapes. *In*: *9th international AAAI conference on web and social media,* Oxford, UK.

Rayson, P., *et al*., 2017. A deeply annotated testbed for geographical text analysis: the corpus of Lake District writing. *In*: *Proceedings of the 1st ACM SIGSPATIAL Workshop on Geospatial humanities*, Redondo Beach, CA, USA, 9–15.

Rosenfeld, S., 2011. On being heard: a case for paying attention to the historical ear. *The American Historical Review*, 116 (2), 316–334. doi:10.1086/ahr.116.2.316

Rundstrom, R.A., 1995. GIS, indigenous peoples, and epistemological diversity. *Cartography and Geographic Information Systems*, 22 (1), 45–57. doi:10.1559/152304095782540564

Schroeder, H.W., 2007. Place experience, gestalt, and the human-nature relationship. *Journal of Environmental Psychology*, 27 (4), 293–309. doi:10.1016/j.jenvp.2007.07.001

Sexton, J.O., *et al*., 2016. Conservation policy and the measurement of forests. *Nature Climate Change*, 6 (2), 192. doi:10.1038/nclimate2816

Smith, B.R., 1999. *The acoustic world of early modern England: attending to the O-Factor*. Chicago and London: The University of Chicago Press.

Smith, M.M., 2004. Introduction: onward to audible parts. *In*: M.M. Smith, ed. *Hearing history: a reader*. Athens, GA: University of Georgia Press, ix-xxii.

Smith, S.J., 1994. SoundScape. *Area,*26, 232–240.

Taylor, J., Gregory, I., and Donaldson, C., 2018. Combining close and distant reading: a multiscalar analysis of the English Lake District's historical soundscape. *International Journal of Humanities and Arts Computing*, 12 (2), 163–182. doi:10.3366/ijhac.2018.0220

Taylor, J.E., 2018. Echoes in the mountains: theromantic Lake District's soundscape. *Studies in Romanticism*, (forthcoming).

Truax, B., 1978. *Handbook for acoustic ecology*. Burnaby, B.C. Canada: ARC Publications.

Tudor, C., 2014. *An approach to landscape character assessment*. (October), 56. Worcester, UK: Natural England.

Turner, S., 2006. Historic landscape characterisation: a landscape archaeology for research, management and planning. *Landscape Research*, 31 (4), 385–398. doi:10.1080/01426390601004376

Van Den Berghe, H., *et al*., 2018. Using the past to indicate the possible presence of relics in the present-day landscape: the western front of the great war in Belgium. *Landscape Research*, 6397, 1–23. doi:10.1080/01426397.2017.1415315

Vasconcelos, M., *et al*., 2002. Land cover change in two protected areas of Guinea-Bissau (1956–1998). *Applied Geography*, 22 (2), 139–156. doi:10.1016/S0143-6228(02)00005-X

Winter, S. and Freksa, C., 2012. Approaching the notion of place by contrast. *Journal of Spatial Information Science*, 2012 (5), 31–50.